Dynamic Mechanism Design for Markets with Strategic Resources

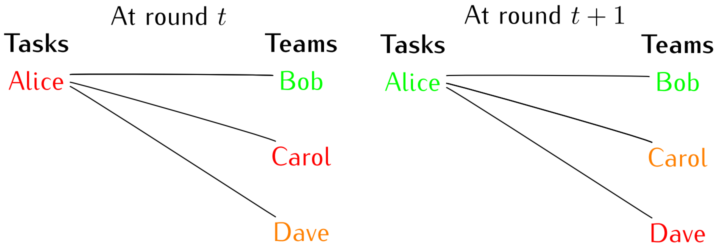
Swaprava Nath[1] Onno Zoeter[2] Yadati Narahari[1] Chris Dance[2]

[1] Indian Institute of Science, Bangalore [2] Xerox Research Centre Europe

Conference on Uncertainty in Artificial Intelligence

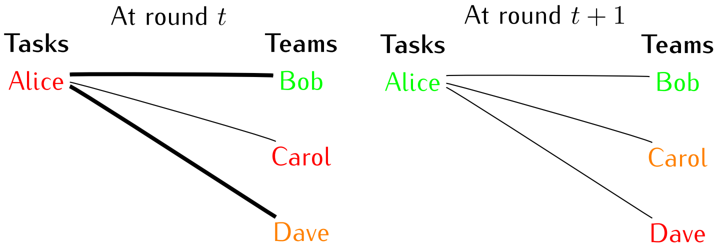
July 15, 2011

Markets with Strategic Resources: An Example



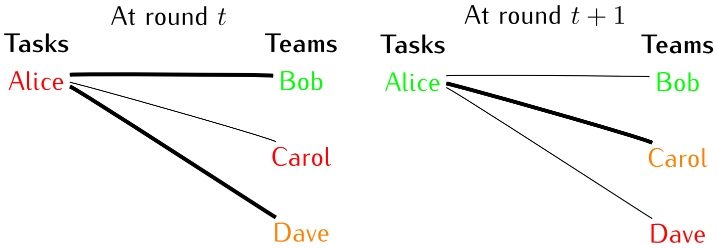
- Task difficulties (Low,Medium,High) and team efficiencies (Low,Medium,High) follow Markov chain.
- Task for central planner: assign teams to tasks in each round, balancing completed task rewards, costs, and future efficiency levels.
- If difficulties and efficiencies are known to planner, this is a Markov decision problem (MDP).

Markets with Strategic Resources: An Example



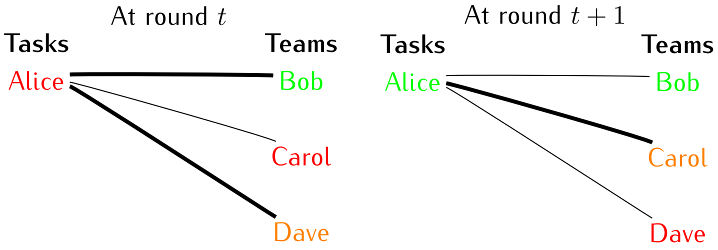
- Task difficulties (Low,Medium,High) and team efficiencies (Low,Medium,High) follow Markov chain.
- Task for central planner: assign teams to tasks in each round, balancing completed task rewards, costs, and future efficiency levels.
- If difficulties and efficiencies are known to planner, this is a Markov decision problem (MDP).

Markets with Strategic Resources: An Example



- Task difficulties (Low,Medium,High) and team efficiencies (Low,Medium,High) follow Markov chain.
- Task for central planner: assign teams to tasks in each round, balancing completed task rewards, costs, and future efficiency levels.
- If difficulties and efficiencies are known to planner, this is a Markov decision problem (MDP).

Markets with Strategic Resources: An Example



- Task difficulties (Low,Medium,High) and team efficiencies (Low,Medium,High) follow Markov chain.
- Task for central planner: assign teams to tasks in each round, balancing completed task rewards, costs, and future efficiency levels.
- If difficulties and efficiencies are known to planner, this is a Markov decision problem (MDP).

This talk: task difficulties and team efficiencies (elements in the state of the MDP) are *private information* of *strategic agents*: this is a *mechanism design* problem.

MDP notation for our setting

Agents 0 task owner (one for ease of notation)

$\{1, \ldots, n\} =: N$ set of resources.

State is concatenation of agent *types*

$$\begin{aligned} \theta_t &= (\theta_{0,t}, \theta_{1,t}, \ldots, \theta_{n,t}) \\ \theta_{i,t} &: \text{ } i\text{'s type, e.g. } \theta_{i,t} \in \{L, M, H\} \end{aligned}$$

Action is an assignment of resources to tasks

$$a_t \in 2^N \text{ for our 1 task example.}$$

State transition function is Markov and **independent** per agent

$$F(\theta_{t+1} | a_t, \theta_t) = \prod_i F_i(\theta_{i,t+1} | a_t, \theta_{i,t}) \ .$$

MDP notation for our setting: interdependent valuations

Reward function is sum of all agents' *valuations* (*social welfare*)

$$\begin{aligned} R(\theta_t, a_t) &= \sum_{i=0}^{n} v_i(a_t, \theta_t) \text{ with} \\ v_0(a_t, \theta_t) &\geq 0 \text{ denoting returns} \\ v_i(a_t, \theta_t) &\leq 0 \text{ for } i > 0 \text{ denoting costs.} \end{aligned}$$

Note: valuations are **dependent**, compare with $v_i(a_t, \theta_{\mathbf{i},t})$.

E.g. task owner's return depends on task difficulty *and* team strength.

MDP goals

Consider infinite horizon problem with discount parameter δ.

Controller's goal is to determine and execute optimal (static) policy π^*

$$
\begin{aligned}
W^*(\theta_t) &= \max_a \left[R(a,\theta_t) + \delta \mathbb{E}_{a,\theta_t} W^*(\theta_{t+1})\right] \quad \textit{(maximal social welfare)} \\
\pi^*(\theta_t) &\in \arg\max_a \left[R(a,\theta_t) + \delta \mathbb{E}_{a,\theta_t} W^*(\theta_{t+1})\right] .
\end{aligned}
$$

Markets with strategic resources

In our strategic setting everything remains common knowledge, *except* θ_t,

$$\theta_{i,t} \text{ is only observed by } i.$$

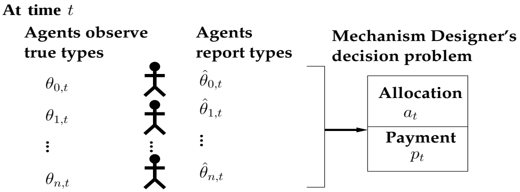
We consider a *quasi-linear setting*: agents care about sum of discounted utilities:

$$\sum_{t=0}^{\infty} \delta^t u_{i,t} \qquad \text{with}$$

$$\begin{aligned} u_{i,t} &= v_{i,t} + p_{i,t} \quad \textit{(utility)} \\ p_{i,t} &> 0 \quad \text{possible } \textit{payment} \text{ from controller to agent} \\ p_{i,t} &< 0 \quad \text{possible payment from agent to controller.} \end{aligned}$$

Mechanism designer's goals

Design a repeated game with information exchange



that achieves

Efficiency (EFF): mechanism yields $W^*(\theta_t)$ under equilibrium reporting strategies.
Truthfulness (Incentive compatibility) (EPIC): it is optimal for i to report $\theta_{i,t}$ truthfully when asked.
Voluntary participation (Individual rationality) (EPIR): agents stand to gain something from participating (non-negative utilities).

We consider (provide proofs for) *ex-post* equilibria: agent i does not make assumptions about other agent's types, but *does* assume that other agents report truthfully.

Strictly speaking *within period* ex-post to emphasizes that agents can't foresee the future.

Where does this work fit in?

Valuations	STATIC	DYNAMIC
Independent	**VCG Mechanism** (Vickery, 1961; Clarke, 1971; Groves, 1973)	**Dynamic Pivot Mechanism** (Bergemann and Välimäki, 2010) (Athey and Segal, 2007) (Cavallo et al., 2006)
Dependent	**Generalized VCG** (Mezzetti, 2004)	

- VCG guarantees
  - DSIC (stronger than EPIC), EFF, under certain conditions EPIR
- GVCG guarantees
  - EPIC, EFF, under certain conditions EPIR
- DPM guarantees
  - EPIC, EFF, EPIR, in non-exchange economies, budget balanced

Where does this work fit in?

Valuations	STATIC	DYNAMIC
Independent	**VCG Mechanism** (Vickery, 1961; Clarke, 1971; Groves, 1973)	**Dynamic Pivot Mechanism** (Bergemann and Välimäki, 2010) (Athey and Segal, 2007) (Cavallo et al., 2006)
Dependent	**Generalized VCG** (Mezzetti, 2004)	**Generalized Dynamic Pivot Mechanism**

- VCG guarantees
  - DSIC (stronger than EPIC), EFF, under certain conditions EPIR
- GVCG guarantees
  - EPIC, EFF, under certain conditions EPIR
- DPM guarantees
  - EPIC, EFF, EPIR, in non-exchange economies, budget balanced
- GDPM guarantees
  - EPIC, EFF, EPIR, but requires more reports from agents than DPM

The Interdependent Value Setting

- If values are dependent, *Efficiency* and *Truthfulness* cannot be guaranteed with single stage mechanisms even in static setting [1]
  - Without imposing any voluntary participation or budget constraints
- Need to split the decisions of allocation and payment [2]

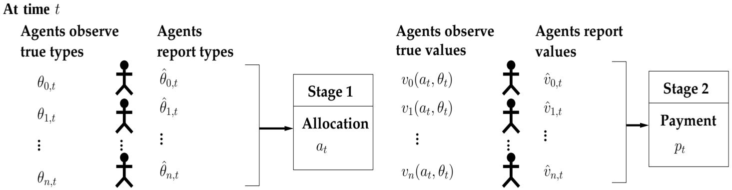
---

[1] P. Jehiel and B. Moldovanu. Efficient Design with Interdependent Valuations. *Econometrica*, (69):1237–1259, 2001.

[2] Claudio Mezzetti. Mechanism Design with Interdependent Valuations: Efficiency. *Econometrica*, 2004.

The Interdependent Value Setting

- If values are dependent, *Efficiency* and *Truthfulness* cannot be guaranteed with single stage mechanisms even in static setting [1]
  - Without imposing any voluntary participation or budget constraints
- Need to split the decisions of allocation and payment [2]



[1] P. Jehiel and B. Moldovanu. Efficient Design with Interdependent Valuations. *Econometrica*, (69):1237–1259, 2001.

[2] Claudio Mezzetti. Mechanism Design with Interdependent Valuations: Efficiency. *Econometrica*, 2004.

The Generalized Dynamic Pivot Mechanism (GDPM)

- The task is to design the allocation and payments on the reported types and values
- The allocation maximizes the social welfare taking reports as truth,

$$a^*(\hat{\theta}_t) \in \arg\max_{a_t} \mathbb{E}_{a_t,\hat{\theta}_t} \left[\sum_{i \in N} v_i(a_t, \hat{\theta}_t) + \delta \mathbb{E}_{\theta_{t+1}|a_t,\hat{\theta}_t} W(\theta_{t+1}) \right]$$

- The payment to agent i at t is given by,

$$p_i^*(\hat{\theta}_t, \hat{v}_t) = \underbrace{\sum_{j \neq i} \hat{v}_{j,t} + \delta \mathbb{E}_{\theta_{t+1}|a^*(\hat{\theta}_t),\hat{\theta}_t} W_{-i}(\theta_{t+1})}_{\text{Expected discounted sum of returns to other agents, based on reported valuations and allocation for this round}} - \underbrace{W_{-i}(\hat{\theta}_t)}_{\text{Const. indep. of } \hat{\theta}_{i,t}}$$

Main Theorem

Theorem

GDPM is efficient, within period ex-post incentive compatible, and within period ex-post individually rational.

Proof ingredients: The allocation and payment is chosen such that

- If everyone reported true $\theta_{i,t}$'s, each would have got their marginal contribution, $W(\theta_t) - W_{-i}(\theta_t)$ as the payoff (check for time instant t).
- Goal: to show that reporting true $\theta_{i,t}$'s maximizes i's payoff, given everyone else is reporting truth (EPIC).
- At t, player i cares about,
  - Current stage payoff, $v_i(a_t, \theta_t) + p_i^*(\hat{\theta}_t, \hat{v}_t)$ and,
  - Future payoffs, i.e., the expected discounted sum of the value + payment from $t+1$ to ∞.
  - From time $t+1$, the expected discounted sum of payoff of agent i is $W(\theta_{t+1}) - W_{-i}(\theta_{t+1})$, assuming agents report truthfully from $t+1$.
- Putting together, agent i's utility is,

$$v_i(a_t, \theta_t) + p_i^*(\hat{\theta}_t, \hat{v}_t) + \mathbb{E}_{\theta_{t+1}|a_t,\theta_t} \left(W(\theta_{t+1}) - W_{-i}(\theta_{t+1}) \right)$$

- This is maximized at the true θ_t reports (proved in paper).

The use of second phase reports

Proof ingredients:

Necessity of the second reporting phase:

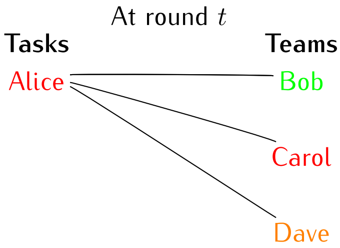
- Controller can only influence assignment.
  With the second reporting phase, i can only influence his payoff via the assignment, i.e. his utility is of a form
  $$f(a^*(\hat{\theta}_{i,t})) \ .$$
  Since controller optimizes what i cares about, truthfulness is optimal.
- Without second phase, payment to i would be based on
  $$v_{j,t}(a_t, \hat{\theta}_{i,t}, \theta_{-i,t}) \quad (j\text{'s predicted value, based on } i\text{'s report}),$$
  instead of
  $$\hat{v}_{j,t} \quad (j\text{'s reported value, which is independent of } i\text{'s report}).$$
  What controller optimizes has form
  $$f(a^*(\hat{\theta}_{i,t}), \hat{\theta}_{i,t}) \ ,$$
  hence i has a richer optimization problem than the controller, and might strategically manipulate his report $\hat{\theta}_{i,t}$.

Why care? A naïve alternative mechanism

Is obtaining efficiency straightforward?

Consider an alternative naïve mechanism

- The allocation maximizes the social welfare taking reports as truth.
- Task owner pays K to every assigned team (independent of outcome).



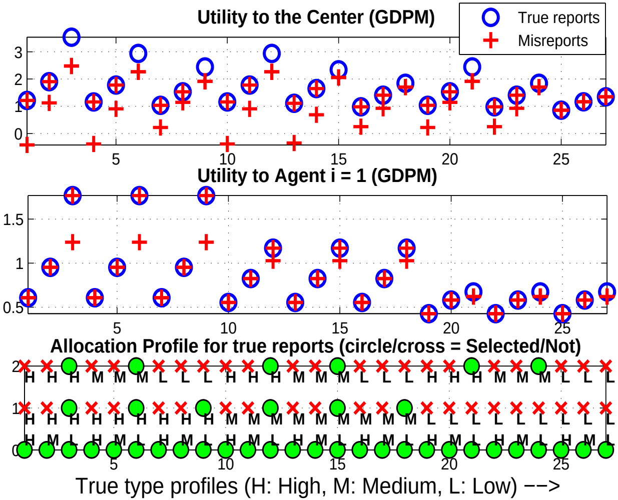
If you were Carol, would you report your low effectiveness state?

Simulation Setting

- 3 players: 1 Task owner (Image owner), 2 Teams (Annotators)
- $\theta_{i,t} \in \{L, M, H\}$ corresponding to the difficulty/effectiveness levels for all agents: $3^3 = 27$ possible states.
- Value structure represents law of diminishing returns.
- Transition probability matrices reflect risk of reduction in effectiveness when assigned, probability of recovery when not assigned.
- Annotators are symmetric, we need to study only one.

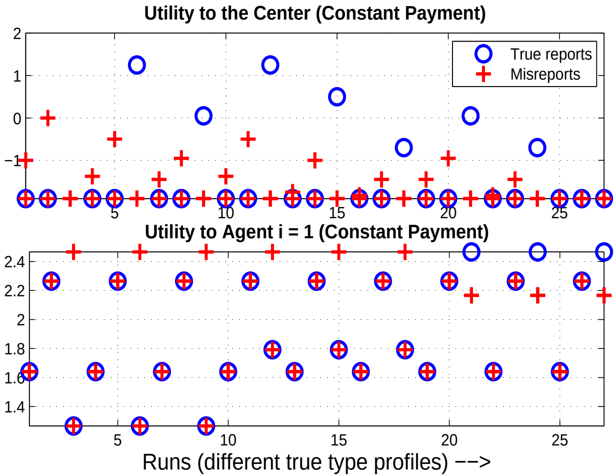
Simulation Results

Truthfulness:



Simulation Results (Contd.)

Comparison with a Naïve Mechanism (CONST):



Simulation Summary

Payment consistency (PC): task owners only make payments, teams only receive payments.
Budget balance (BB): controller does not need to inject money into the exchange.

	EFF	EPIC	EPIR	PC	BB
GDPM	✓	✓	✓	×	×
CONST	×	×	×	✓	✓

- All of these properties may not be simultaneously satisfiable

Discussion

Strategic extensions of dynamic decision problems are very important in practical problems.

We have presented a dynamic mechanism for exchange economies.

It is (within period, ex-post)
truthful, efficient, and voluntary participatory
but not
budget balanced, payment consistent
in a setting with
independent type transitions, and dependent valuations.

See also Cavallo et al. '09 who consider dynamic problems with
dependent type transitions, and independent valuations.

Future work: complete this space and determine (im)possibilities.

What extra opportunities are there in the infinite discounted case over the single round setting?

Questions?